DOCUMENT RESUME

ED 455 312 TM 033 172

AUTHOR Takona, J. ole
TITLE Perspective on Grade Assignment at East Africa's State Universities.
PUB DATE 2001-00-00
NOTE 15p.
PUB TYPE Opinion Papers (120) -- Reports - Descriptive (141)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS College Students; Foreign Countries; *Grades (Scholastic); *Grading; Higher Education; *Student Evaluation
IDENTIFIERS *Africa (East)

ABSTRACT

This paper discusses approaches to grading students that are being used or could be used in state universities in East Africa. Grade assignment usually means that some kind of comparison is being made, whether with other students or established standards or based on improvement and ability. For the most part, university regulations encourage and provide for a uniform grading policy across faculties and departments, although in practice grading may vary by department or discipline and over time. Approaches that may be used include a distribution gap approach, grading on the curve, and a percent grading approach. A grading method that seems appropriate for many African universities with large course enrollments (100 students or more) is that of using group comparisons to provide a reference group representative of students typically enrolled in the course. In a majority of state universities in East Africa, it is not uncommon to find multisectional and team-taught courses. How an examiner assigns course grades in such a situation is an important consideration. To establish common grading components in each course section, agreement among lecturers and with the department should be reached at the beginning of a course. Examiners should compare their grade distributions with grade distributions for similar courses in the same department. (Contains 15 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 455 312

Perspective on Grade Assignment at East Africa's State Universities

J. ole Takona, Ph.D.

TM033172

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

J. Takona

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

2

PERSPECTIVE ON GRADE ASSIGNMENT AT EAST AFRICA'S STATE UNIVERSITIES

J. ole Takona, Ph.D.

Grading is an integral part of university education. Grades assigned by examiners are intended to convey a level of achievement by student in the course. Unless uniform standards in assigning grades are used, assigned grades are apt to misinform and mislead those who use them. With careful thought and periodic review, university examiners can develop satisfactory and defensive grading policies and procedures. Hess, et al (1987) and Vasa (1981) indicated that the position of pre-eminence given to grading becomes most apparent when considered in light of the varied functions that grades serve for administrators, students, teachers, guidance counselors, and parents. Hess, et al. noted that "grading is an important aspect of documenting the educational experience of students [and thus] the assignment of grades has created and will continue to create debate within the educational community" (p. 1).

Grading has remained a controversial topic in Education with hardly any research studies undertaken to establish sound policies and practices. In general, the literature on grading has included position papers addressing grading policies and practices rather than empirical studies (e.g., Carpenter, 1985; Ornstein, 1989; Terwilliger, 1977). Although the integral role of grading is often stressed, its policies and practices have been criticized as being too simplistic, overlooking more intangible outcomes of learning, being ineffective motivators, failing to emphasize mastery as the foremost concern, leading to failure, subjecting all students to universal standards, and promoting competition among students rather than cooperation (Terwilliger, 1977).

There are no research studies to answer questions like: What should an "A" grade mean? What percent of the students should receive a "C?" Should writing skills, creativity, or clarity of arguments be judged in assigning a grade or should content alone remain as the basis? What should a course grade represent? These questions require value judgments rather than an interpretation of research data; the answer to each may vary from examiner to examiner but they must find acceptable answers in order to establish defensible standards of grading policies and practice. Fair and explicit grading policies are an important aspect of quality instruction.

It is not sufficient to have a standardized method if it is not defensible in terms of the goals of university education and the realities of the setting in which grades are given. This being so, an examiner's view of the role of university education affects his grading plans. One who holds that the end product of a university education should be a "prestigious" group which has survived sixteen or more years of culling and sorting may hold a different grading policies from the one who believes that most students who qualify for university entrance in the first place are capable of earning a degree within four years with an acceptable grade.

A number of factors influence philosophical beliefs to instruction. As any of these factors change a corresponding change in grading often occurs. Professional educators have developed a variety of models of instruction, each designed to produce classroom learning. However, each model differs in the specific type or measure of learning that is targeted. Therefore, as we make decisions about "best educational practices" we must be certain that we connect recommended prac-

tices with desired outcomes. For example, one instructional strategy may dictate the type of grading procedures to use. Consider for a minute the mastery learning approach (Block, 1971) to teaching which would require the reproduction of content material. Another may be based on command of general concepts and not necessarily on their reproduction. For instance, an application of an appropriate statistical procedure to a set of data may carry more weight than the obtained answer in a Research Methodology course. While yet another grading system may be based on an arbitrarily set number of "A" or "B" grades.

University grading policies do, in fact, limit the procedures that can be used, thereby, forcing a basic grading plan on examiners. The recent response to grade inflation in American universities (Cross, 1993; Bassinger, 1997; Smith, 1998) might cause some examiners, individually and collectively, to alter their philosophies and procedures. Further pressure from external moderators in African universities to give lower or higher grades often causes some examiners to operate in conflict with their own views.

Where students' evaluation of instruction is undertaken, the need for positive evaluation may contribute to the shaping or altering of the grading philosophies of some examiners. On the other hand, the dissonance created by institutional regulations probably contributes to the wide-spread feeling that end-of-course grading is one of the least pleasant tasks.

COURSE GRADE ASSIGNMENT APPROACHES

Grade assignment, essentially means that some kind of comparison is being made. There are a number of criteria to which an examiner may reference. Here we will discuss a few common comparisons used to determine course grades and point out the merits and demerits associated with each.

Comparisons with other students

By comparing a student's overall course performance with that of a reference group, an examiner assigns a grade depicting a student's level of achievement within that cohort. The reference group may include all students enrolled in the course during a given semester in the past or all students enrolled in a course since its inception. This comparison group holds the key to present and future interpretation of student performance. An "A" might not represent excellence in attainment of knowledge and skill if the reference group as a whole is somewhat inept.

Two primary advantages are inherent in such an approach. Firstly, students whose academic performance is outstanding in comparison to their peers are rewarded. Secondly, The system is a common one that many examiners are familiar with. Given additional information about the students, examiners, or department, grades can be interpreted easily.

The disadvantages of this approach are: no matter how outstanding the reference group of students is, some will receive low grades; no matter how low the overall achievement in the reference group, some students will receive high grades. Further, grades are difficult to interpret without additional information about the overall quality of the group and degree of difficulty of the overall examination.

Further, grading standards in a course may fluctuate with the quality and composition of each class of students. The performance of a bright class raises the standard and a less able group of students would lower it. Often a student's grade depends on the composition of the cohort. Third, there exists a need to develop course "norms" which account for more than a single class performance. This requires that the examiner has access to a grade record of several previous cohorts. An examiner who is new to the course may place the students at a particular disadvantage since the reference group will necessarily be small and very possibly atypical.

Comparisons with established absolute standards

Grades may be assigned by comparing a student's performance with a set of specified absolute standards rather than relative standards such as the work of other students. In this grading method, the examiner is interested in indicating how much of a set of tasks or ideas a student knows, rather than how many other students have mastered more or less of that domain. Such absolute standards are often spelled out in the examination marking schemes.

Ideally, student's grades depend solely on their attained level of content mastery; thus the levels of performance of their classmates has no bearing on the final course grade. There are no quotas in each grade category. It is possible in a given class that all students could receive an "A" or a "B" or any other grade. A "C" in an introductory statistics class might indicate that the student has minimal knowledge of descriptive and inferential statistics. A much higher achievement level would be required for an "A."

There are a number of advantages associated with the approach: First, course objectives and standards must be clearly defined and communicated to students in advance. Further, with an effort on the part of the student coupled with adequate instruction, higher grades can be obtained. Emphasis, however, should be focused is on achieving course objectives, rather than competing for a grade.

On the contrary, it is difficult and time consuming to determine what the standards should be for each possible course grade issued. Inexperienced examiners may be at a disadvantage in making these assessments. It is useful, however, to note that a sound interpretation of the meaning of a course grade cannot be approximated accurately unless course objectives are clearly defined. These must be provided in a course syllabus and serves as a guide for each student and, therefore, must be presented at the beginning of the course. In addition, such an instrument must provide detailed information with respect to course requirements, student expectations, schedule of activities, dates for Continuous Assessment Tests (CATs) and examinations, paper deadlines, criteria for evaluating performance, and tips on how to succeed. The examiner has to decide on reasonable expectations of students and necessary prerequisite knowledge for subsequent courses.

Comparisons based on improvement and ability

Improvement and ability are sometimes used in grading students. Grades may be based on students' exit behavior when compared to their entry behavior as measured at the beginning of the

course. Large gains are assigned high grades and small gains are represented by low grades. With this method, a student who enter a course with some pre-course knowledge have less to gain from a course than does a relatively naive student. From a measurement stand-point, the post test-pretest gain score is more error-laden.

Although growth is certainly important when assessing the impact of instruction, it is less useful as a basis for determining course grades than end-of-programme competence.

Generally, students with high entry scores in their O-Level (e.g. Kenya Certificate of Education) would be expected to achieve higher course grades than those with lower incoming scores. When grades are based on comparisons with predicted ability, an "overachiever" and an "underachiever" may receive the same grade in a particular course, yet their levels of competence with respect to the course content may be vastly different. The first student may not be prepared to take a more advanced course, but the second student may be. A grade might represent the amount students learned in a course relative to how much they could be expected to learn as predicted from their measured academic ability. A grade may, in part, reflect the amount of effort a student has put into a course. The high ability students who can satisfy course requirements with minimal effort are penalized for their apparent "lack" of effort. Since the letter grade alone does not communicate such information, the value of ability-based grading does not warrant its use.

Further, serious philosophical and methodological problems related to these comparisons arise making such an approach highly questionable for most educational situations within the Kenya's state university system where applicants' aggregate scores in KCE lack significant variability.

A single course grade should represent only one of the several grading comparisons noted above. To expect a course grade to reflect more than one of these comparisons is too much of a communication burden. Examiners, who wish to communicate more than relative group standing, or subject matter competence or level of effort, must find additional ways to provide such information on each student.

For the most part, university's regulations encourage and provide for a uniform grading policy across Faculties and Departments. In practice, grade distributions vary by department, by Faculty and discipline and over time within each of these units (Takona 1999). Other units usually know the grading standards of a Department or Faculty. For example, a "B-" in a required course given by Department X might indicate that the student probably is not a qualified candidate for post-graduate programmes or a related field. Or, a "C+" in a required course given by Department Y might indicate that the student's knowledge is probably adequate for the next course. As a consequence, an examiner may provide grades that may be misleading about a student and or may also be misinterpreted.

If grading pattern of an examiner differs markedly from others in the department the examiner should re-examine his or her grading practices to see that they are rational and defensible. In some instances an individual examiner's grading practice may differ markedly the department's and yet be defensible. For example, an examiner may be using different grading standards, or a course structure grading plan which differs from departmental guidelines, or his ideas about the

function of grading may differ from the department's. In such cases, a satisfactory grading plan can be worked out.

Following successful instruction, students must demonstrate their achievement of learning by providing a product validating that achievement. Carefully written graded assignments and examination are key to accurate validation. Because it is not customary at the university level as it is in primary school to accumulate many grades per student, each grade carries great weight and should be as accurate as possible. Poorly planned CATs and assignments increase the likelihood of basing grades primarily on factors of chance. And it should, therefore, not be surprising that some would argue that over the course of a university education, students would receive an equal number of higher-grades-than-merited and lower-grades-than-merited. Consequently, final GPA will be relatively correct. However, in view of the many uses of grades, each course grade is significant in itself to the student and others.

At Moi University, lecturers often recognize their responsibility to communicate to students the manner in which grades will be determined. A good course syllabus often indicates the number and types of tests to be administered and how much each test, homework assignment, and other course requirements will count towards the final course grade. The weight of these activities is dictated by the University's regulations, which expects that they compose 30 percent of the final grade. However, it is up to the examiner to determine the weighting to be assigned to each individual activity.

By stating the grading procedures at the beginning of a course, lecturers are essentially making a "contract" with the class about how each student will be evaluated. This being the case, the contract should provide the students with a clear understanding of the expectations so they can structure their work efforts.

Students, however, have a right to know what method will be used in to assigning a final grade and what kind of comparison the course grade will represent. By informing students early in the semester about course priorities, they are encouraged to study what is deemed valuable. An indication must, therefore, be stipulated on: which course activities will be considered in course grade; weight of CATs, term-papers and final paper; and which topics are more important than others.

Two common complaints by students at any given institution are that grading procedures stated at the beginning of the course were either inconsistently followed or were changed without explanation or even advance notice. Altering or inconsistently following the grading plan is analogous to playing a game wherein the rules arbitrarily change, sometimes without the players' knowledge. Such a scenario diminishes students' ability to participate placing students in the unreasonable position of never knowing for sure what is expected of them. For this reason, the grading system must be reliable and consistent.

From a decision-making point of view (Drucker, 1999), the more pieces of information available to the decision-maker, the more confidence they will be in forming an accurate and appropriate decision. This same principle applies to the process of assigning grades. If only a final exam score is used to assign a course grade, the adequacy of the grade will depend on how well the test

covered the relevant aspects of course content and how typically the student performed on one specific day during a 2-3 hour period.

Though the minimum number of CATs, term-papers, projects, and other assignments needed must be course- specific, each examiner must attempt to secure as much relevant data as is reasonably possible to insure an accurate reflection of each student's achievement level.

Even at a university where grading regulations are standard, various grading practices are used. Examiners often wish to weight some components more heavily than others do. For example, two combined CAT marks may be valued at the same weight as one term paper. When assigning weights consideration ought to be given to the extent to which each grading component measures important goals; what level of achievement can be accurately measured with each grading component; and most importantly, that each graded component measures a specific course content or objectives.

Once the weight of each grading component has been established, the examiner must insure that the desired weights are actually used. This task is not as simple as it first appears. An extreme example of weighting will illustrate the problem.

Suppose an administration of a 50-item CAT and a 100-item CAT is to be combined so they have equal weight (50 /50 percent in the aggregate). We must know something about the spread of scores or variability (e.g., standard deviation) on each CAT before adding the scores together. Assume that scores on the shorter CAT are quite evenly spread throughout the range 20-50, and the scores on the other are in the range 90-100. Because variability on the 100-item CAT is so small, if we merely add each student's scores together, the second CAT will have very little weight in the total score and the spread of scores in the aggregate be emulate the spread of scores observed on the first CAT. The net effect is analogous to adding a constant value to each student's score on the 50-item CAT the students maintain essentially the same relative standing. Figure 1 demonstrates how scores can be adjusted to achieve the desired weighting before combining them. CAT No. 2 is twice as long as the first, but there is twice as much variability-ability in CAT No. 1 scores. This is the "observed weight."

The standard deviation approximates the extend to which scores deviate from the mean. The larger the value, the more the scores are spread across the possible range. The variability of scores, therefore, is the key to proper weighting. If we merely add these scores together, CAT No. 1 will carry 66 percent of the weight while Exam No. 2 will carry 33 percent weight. We must adjust the scores on the second CAT so that the standard deviation of the scores will be similar to that for CAT No. 1 by multiplying each score on the 100-item CAT by two; the adjusted scores will become more varied (standard deviation = 9.0). The score from CAT No. 1 can then be added to the adjusted score from CAT No. 2 to yield a total in which the components are equally weighted. A practical solution to combining several weighted components is to first transform raw scores to standard scores, z or T, before applying relative weights and adding (Ebel & Frisbie, 1991; Linn & Gronlund, 1995; and Ory & Ryan, 1993). After grading weights have been assigned and combined, scores are calculated for each student.

Distribution Gap Approach

Another approach to course grade assignment is the distribution gap approach which is based on the relative ranking of students in the form of a frequency distribution their scores. The frequency distribution is carefully scrutinized for observed gaps with several consecutive scores which have zero frequency. A horizontal line is drawn at the top of the first gap indicating the cut-off of "A's" and a second gap is sought for the next grade. The process continues until all possible grade ranges (A-F) are identified. The major fallacy with this technique is its dependence on "chance" to form the gaps. The gaps are random because measurement errors that my include guessing, poorly written items, dictate where gaps will or will not appear. If scores from an equivalent test could be obtained from the same group, the gaps would likely appear in different places. Some students would get higher grades, some would get lower grades, and many grades would remain unchanged. Furthermore, unless the examiner has additional achievement data to re-evaluate borderline cases, many students could see their fate determined more by chance than performance.

Grading on the Curve Approach

Examiners are often faced with the familiar question: "How come I got a C+?" Most students have difficulty understanding the mysterious process by which their hard work is reduced to a single mark. It is here that grading on the curve becomes handy. This method of assigning grades based on group comparisons is complicated by the need to establish arbitrary quotas for each grade category. The question that now arises is what percent should get A's? B's? D's? Once these quotas are fixed, grades are assigned without regard to level of performance. The highest 10 percent may have achieved at about the same level. Those who "set the curve" are merely among the top group. Their grade may be the same as that of a student who scored 20 points lower. The bottom five (5) percent may be assigned "F's" though the bottom 15 percent may be relatively indistinguishable in achievement. Quota-setting strategies vary from examiner to examiner and department to department and seldom carry a defensible rationale.

While defense for the use of the familiar bell shaped curve as an appropriate model for setting quotas can be made, using the normal curve is as arbitrary as using any other curve. It is highly unlikely that university student abilities or achievement are normally distributed, especially in Kenya's state universities where admittance depends on performance on the national examination and the number of vacancies available. If vacancy allows an admission of only 6,000 students and 6,000 students with an average of an A in KCE are found, then only those who achieved an A will be admitted. While grading on the curve is efficient, its usefulness in an African university may be far fetched.

Percent Grading Approach

Various score components are typically converted to a percent based on the total possible score. The percent score is then interpreted as the percent of content, skills or knowledge over which the student has command. Thus an examination score of 83 percent means that the student knows 83 percent of the content which is represented by the test items.

The long-standing use of percent grading in any form is questionable. In percentage grading, an examiner establishes that arbitrary standards similar to those set for grading on the curve, i.e. when grades are assigned a percentage score, students with scores 93 - 100 get A's and 85-92 is a B, 78-84 is a C, etc. The restriction here is on the score ranges rather than on the number of individuals who can earn each grade. Should the cutoff for an A be 92 instead? Why not 90? What sound rationale can be given for any particular cut-off? In addition, it seems indefensible in most cases to set grade cut-offs that remain constant through- out the course and several consecutive offerings of the course. It does seem defensible for the examiner to decide on cutoffs for each grading component, independent of the others, so that the scale for an A might be 93 - 100 for CAT. 1, 88 - 100 for a paper, 87-100 for CAT . 2 and 90 -100 for the final paper.

A problem, however, does occur when the highest score obtained on an examination is only 68 percent, for example. Was the examination much too difficult? Did students study too little? Was instruction relatively ineffective? Oftentimes, examiners decide to "adjust" scores using a factor so that 68 percent, in this instance, is equated to 100 percent. Though the adjustment might cause all concerned to breathe easier, the new score is essentially the percentage of examination paper`s content learned by the students. The examination score of 83 no longer means that the student knew 83 percent of its content.

A grading method that seems appropriate for many African universities with large (100 or more) course enrollments is that using group comparisons to provide a reference group representative of students typically enrolled in the course. The following steps describe a widely-used and generally sound procedure:

- Convert raw scores on each exam to a standard score (z or T) by using the mean and standard deviation for each graded component. Use of scores is adequate because they allow an approximation of performance on each grading component with an identical yardstick. When relative comparisons are to be made, it is not advisable to convert raw scores to grades and average the separate grades. This is because the distinction between achievement levels will be lost; differences will melt together as students are forced into a few broad categories.

- Weight each grading variable before combining the standard scores. For example, if the graded components include two CATs, a term-paper, a group project, and a final examination, double both standard scores the two CATs and the term-paper, do nothing to the standard score for the group project and triple the final exam standard score. The respective weights for these variables in the total will then be 20 percent, 20 percent, 20 percent, 10 percent, and 30 percent.

- Add these weighted scores to get a composite score.

- Build a frequency distribution of the total scores by listing all obtainable scores and the number of students receiving each. Calculate the mean, median, and standard deviation.

- If the mean and median are similar in value, use the mean for further computations. Otherwise use the median. Let's assume we have chosen the median. Add one half of the standard deviation to the median and subtract the same value from the median. These are the cutoff points for the range of C's.

- Add one standard deviation to the upper cut-off of the C's to find the A-B cut-off. Subtract the same value from the lower cutoff of the C's to find the D - F cut-off.

Use the number of assignments completed or quality of assignments or other relative achievement data available to re-evaluate borderline cases. Measurement error exists in composite scores too!

Although university regulations are a standard benchmark for grade cut-offs, it should be the examiner's responsibility to decide logically on the values to be used for finding grade cut-offs (one-third, or three-fourths of a standard deviation, for example). How the current class compares to past classes in ability should be judged in setting standards. When B rather than C is considered the average grade, step five will identify the A-B and B-C cut-offs and six would be changed accordingly.

Relative grading methods like the one outlined above are not free from limitations; subjectivity enters into several aspects of the process. But such a systematic approach is not likely to be subject to charges of capricious grading and miscommunication.

Absolute Standard Grading Approach

A prevalent approach at Moi University is the Absolute Standard Grading Method which requires a description of the expected learner's behavior at the end of instructional intervention so that grading components can be determined and measures can be built to evaluate performance. Instructional objectives ought to are provided a guide to students of their learning thus enabling them to related to achievement measures. In principle, each time achievement is measured, the score is compared with some criterion provided in a comprehensive marking scheme. Students who do not meet the minimum criterion level study further and be evaluated again in a supplementary examination.

VALID GRADING COMPONENTS

A distinction should be made between components which a lecturer evaluates and components which are used to determine course grades. Components which contribute to determining course grades should reflect each student's competence in the course content. The components of a grade should be academically oriented and should not be tools of discipline or awards for pleasant personalities or "good" attitudes. It is natural to get attracted to students who are agreeable, friendly, industrious, and kind; and be repelled by those with opposite characteristics. To the extent that certain personalities may interfere with class work or have limited chances for employment in their field of interest, constructive feedback from the lecturer may be necessary. While it might be a good idea for examiners to give feedback A student's personality should, however, not have a direct bearing on course grading.

An "A" in a course should indicate that the student has a firm grasp of the skills and knowledge taught in that course. If the student is merely marginal academically but very industrious and congenial, an "A" grade would be misleading and would diminish the motivation of the excellent students enrolled in the same program.

Grading Attendance and Participation

Attendance and class participation is a potential component whose validity for inclusion in a grading plan is defensible. University students should be encouraged to attend class meetings because it is assumed that the lectures, demonstrations, and discussion will facilitate their learning. At Moi University, regulations dictate that students missing more than 20% of scheduled lecture sessions must withdraw their enrollment. This policy has rarely been enforced partly because of the difficulty and the cost of taking roll calls in lecture rooms with students often exceeding 300. However, it should be emphasized here that students who miss several lecture sessions are more likely to cause a direct detrimental effect on their performance on the course. If the lecturer further reduces the course grade because of absence, the lecturer is essentially submitting such students to "double jeopardy." For example, a lecturer may state that attendance counts toward a given percentage of the course grade. Caution must, however, be taken here to make sure that such policies will not violate university's grade mark regulations. Lecturer who experience a high level of "cutting" might do well if they examine their classroom environment and instructional methods to determine if changes are needed and ask their students why attendance was low.

Obviously seminars and small classes, and especially in the post-graduate programmes, depend on student participation to some degree for their success. When participation is important, it may be appropriate for the lecturer to use participation grades. In such cases the lecturer should keep weekly notes regarding frequency and quality of participation; waiting until the end of the semester and relying strictly on memory makes a relatively subjective task even more subjective. Care must be taken if participation is graded in most courses due to the fact that dominating or extroverted students tend to win.

Grading Workmanship

Neatness in written work, correctness in spelling and grammar, and organizational ability, creativity and originality are all worthy traits. To this extent it seems appropriate that examiners include these factors in their evaluation and give students feedback about them.

Grading Multi-Sectional Courses

In a majority of state universities in East Africa, it is not uncommon to find multi-sectional and team-taught courses. At Moi University's Department of Educational Communication and Technology, teams of up to a dozen lecturers teach most courses. How would examiner's get around the issue of assigning course grades is such a situation?

Unique grading problems are associated with large multiple-sectioned courses taught by different lecturers under the direction and leadership of one course coordinator. In many of these situations there is a common course outline or syllabus, common text, and a set of common classroom tests. The course coordinator is often concerned about the potential lack of equity in grading standards and practices across the many sections. To promote fairness and equality, the following conditions might be established as part of course planning and monitored throughout the semester by the course coordinator:

The number and type of grading components should be the same for each section. This being so the following conditions ought to provide guidance to those involved in assessing student learning:

- All grading components should be identical or nearly equivalent in terms of content measured and level of difficulty.

- Section lecturers should agree on the grading standards to be used (e.g., cut-off scores for graded components; weights to be used with each component in formulating final course score; and the level of difficulty of test questions to be used).

- Method of assigning scores to essays, papers, lab write-ups, and presentations and others should be consistent across sections.

Although all of these conditions can be addressed in the course planning stage, their implementation may be a difficult task. Successful implementation requires a spirit of compromise between section lecturer and the course coordinator as well as among section lecturers. Frequent review of lecturer's practices by the course coordinator or Dean and constructive feedback to the lecturers are necessary. The following additional guidelines contain suggestions for promoting equity in grading across multiple sections:

- To establish common grading components in each course section, an agreement should be reached at the beginning of the course on the components to be used. Agreement should also be reached on the component weighting scheme and final requirements for each course grade.

- To encourage instructional adequacy across sections, course coordinators must distribute the same course objectives, outlines, lecture notes and handouts to all section lecturers. If each lecturer is allowed to contribute to the construction of common tests, CATs, or projects, the section lecturers will become more aware of important course content and the expectations of the course coordinator. This awareness will serve to "standardize" section lectures, also.

- Prior to the administration of an exam, CAT or project, all lecturers should agree on established letter grade cutoff scores. Such consensus helps standardize the administration of grading procedures by reducing the number of examiners who wish not to conform to someone else's standards.

In cases where the grading of particular components is more subjective, organized group practice helps to unify the application of evaluation procedures. For example, course coordinator may wish to distribute examples of A, B, or C quality projects to section lecturer as models prior to the grading of their own class projects. Or, groups of examiners may wish to practice grading a stack of essay exams by circulating and discussing their individual ratings. Through such group practice those involved can compare their evaluation practices with one another thus becoming more uniform over time.

Any grading changes made in a particular section should be implemented in all sections. Examiners can compare their grade distributions with the grade distributions for similar courses in the same department.

To illustrate this: An examined section of a 200-level course with 40 students yields a grade distribution turned out to be: A = 2% B = 30% C = 46% D = 19% F = 3%. When this grade distribution is compared with all of the previous year's sections of the same course, you found the following grade distribution: A = 15% B = 20% C = 35% D = 25% F = 5%. Because your grade distribution is not significantly consistent with departmental records, further investigation is warranted to find out if this particular class was atypical; if examination expectations were too high; if the examinations upon which the grades were based were too difficult and so forth. The fact that the grade distribution does not resemble the grades assigned in the previous year does not necessarily indicate that the grading method was incorrect or inappropriate.

However, discrepancies that you regard as significant should suggest the need for re-examination of your grading practices in light of departmental and general university regulations. On the other hand, it might be a good idea to heed to the advise of theorists and researchers (Ames and Ames, 1991) when they advocate a need to focus on helping students adopt a healthy attitude toward grades by emphasizing individual progress rather than comparing students.

REFERENCES

Ames, R., and C. Ames (1991). "Motivation and Effective Teaching." In L . Idol and B. F. Jones (eds.) Educational Values and Cognitive Instruction: Implications for Reform. New York: Lawrence Erlbaum.

Bassinger, D. (1997). "Fighting Grade Inflation: A Misguided Effort?" College Teaching. 45: 88-91.

Block, J. H. (1971). Mastery Learning; Theory and Practice. New York: Holt, Rinehart and Winston.

Carpenter, D. (1985). Grading Handicapped Pupils: Review and Position Statement. Remedial and Special Education, 6(4): 54-59.

Cross, T. L. (1993). "On Scape-goating Blacks for Grade Inflation." The Journal of Blacks in Higher Education.

Drucker, P. F. (1999). Management Challenges for the 21st Century. San Francisco, CA: Harper Collins Publishers.

Ebel, R. L., & Frisbie, D. A. (1991). Essentials of Educational Measurements (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Hess, R., Miller, A., Reese, J., & Robinson, GA (1987). Grading-Credit-Diploma: Accommodation Practices for Students with Mild Disabilities. Des Moines, IA: Department of Education. (ERIC Document Reproduction Service No. 294 403).

Linn, R., & Gronlund, N. (1995). Measurement and Assessment in Teaching. (7th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Ornstein, A. C. (1989). "The Nature of Grading." The Clearing House, 62: 365-369.

Ory, J.O. and Ryan, K. E. (1993). Writing and Grading Classroom Examinations. Newbury Park, Ca: Sage.

Smith, R. A. (1998). "The Crisis in Cultural Values. Wither Policy for Arts Education." Arts Education. 99: 2-5.

Takona, J. P. ole, (1999). Distribution of Undergraduate Examination Questoins among the Specified Cognitive Levels: A Case of an African University. (ERIC Document Reproduction Service ED 444429).

Terwilliger, J. S. (1977). "Assigning Grades: Philosophical Issues and Practical Recommendations." Journal of Research and Development in Education. 10(3): 21-39.

Vasa, S. F. (1981). "Alternative Procedure for Grading Handicapped Students in Secondary Schools." Education Unlimited. 3: 16-23.